FAKULTÄT
FÜR !NFORMATIK
Faculty of Informatics

Getting Insight on Animal Behaviour through Interactive Visualization of Multiple T-Maze Ensembles.

Das Gate-O-Gon

BACHELORARBEIT

zur Erlangung des akademischen Grades

Bachelor of Science

im Rahmen des Studiums

Medieninformatik und Visual Computing

eingereicht von

Fabrizia Bechtold
Matrikelnummer 00827420

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Privatdoz. Dipl.-Ing. Dr. techn. Krešimir Matković

Wien, 20. August 2020

Fabrizia Bechtold	Krešimir Matković

FAKULTÄT
FÜR !NFORMATIK

Faculty of Informatics

Getting Insight on Animal Behaviour through Interactive Visualization of Multiple T-Maze Ensembles.

The Gate-O-Gon

BACHELOR'S THESIS

submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science

in

Media Informatics and Visual Computing

by

Fabrizia Bechtold
Registration Number 00827420

to the Faculty of Informatics

at the TU Wien

Advisor: Privatdoz. Dipl.-Ing. Dr. techn. Krešimir Matković

Vienna, 20th August, 2020

Fabrizia Bechtold

Krešimir Matković

Danksagung

Ein besonderer Dank gehört dem Team des VRVis Research Centers, im Speziellen Privatdoz. Dipl.-Ing. Dr. techn. Krešimir Matković, für die Möglichkeit mit Ihnen diese Arbeit zu schreiben, für Ideen und anregende Diskussionen, und vielen Opportunitäten, die sie mir geboten haben.
Ich möchte mich auch bei meiner Familie bedanken für ihre Unterstützung und Hilfe wärhend der Erarbeitung dieser Bachelor These und meines gesamten Studiums.

Acknowledgements

A special thanks belongs to the whole team at the VRVis Research Center — especially Privatdoz. Dipl.-Ing. Dr. techn. Krešimir Matković — for providing the opportunity of doing this thesis together with them, for the ideas and stimulating diskussions, and for all the opportunities they provided for me.
Additionally, I want to thank my family for supporting and helping me throughout writing this thesis and my whole studies.

Kurzfassung

Ethologen und Verhaltensforscher erfoschen kognitive Fähigkeiten — wie Lernen und Gedächtnisleistungen — von Nagetieren mit dem Ziel ein besseres Verständnis dafür zu bekommen, wie ähnliche Prozesse in Menschen ablaufen könnten. Solche Studien basieren oft auf Experimenten. Für Experimente, wie dieses welche die Grundlage für diese These bildet, werden Nagetiere mehrere Male in ein Labyrith gesetzt und beobachtet. Dabei wird der Pfad den die Tiere durch das Labyrith nehmen aufgezeichnet und analysiert. Typischerweise erfolgt die Analyse solcher Pfad-Trajektorien auf simplen Eigenschaften und Standardstatistiken, wobei je ein Pfad einzeln untersucht wird. Im Normalfall ist es nicht möglich mehrere Pfade simultan zu vergleichen. Gemeinsam mit Experten auf dem Gebiet der Verhaltensforschung wurde der typische Ablauf einer Analyse von Pfaden aus einem multiblen T-Labyrith in Augenschein genommen und abstrahiert. Basierend darauf wurde ein interaktives, visuelles Analysesystem entwickelt. Das Ziel dieses Systems ist es die Analysearbeit zu vereinfachen und einen tieferen und neuen Einblick auf Lernfähigkeit und Entscheidungsfindung zu gewähren. Zunächst wird ein Überblick über ähnliche Werke und artverwandte Analysesysteme gegeben. Anschließend folgen eine Deklaration der Anforderungen an die Analysearbeiten und eine klare Aufgabenaufschlüsselung. Die Datenerfassung, Datenaufbereitung und -aggregierung wird erläutert und ein Einblick in die zugrunde liegende Datenstruktur gegeben. Das entwickelte System — der T-Maze Explorer — unterstützt multible, verlinkte Ansichten — auch "Views"genannt —, welche etliche Visualisierungstechniken bereitstellen. Zusätzlich werden zwei neue Visualisierungsmethoden vorgestellt, welche im Zuge dieser Arbeit entwickelt wurden. Das T-Maze View zeigt alle Pfade aus einem Analyseensemble im Kontext des Labyriths und gibt zusätzliche Funktionalitäten, wie das Hervorheben von richtungswechselnden Pfaden. Das Gate-O-Gon View extrahiert aus den Pfaden wie oft ein Tier auf dem Weg durch das Labyrith umgekehrt ist, und in welchen Teilen des Labyriths Umkehrungen stattfinden. Diese Information wird in einem kompakten und informationsdichten Element dargestellt. Der T-Maze Explorer hat es zum Ziel seinen Anwendern das Auffinden von möglichen Mustern beziehungsweise abnormalen sVerhalten in der Bewegung der Tiere zu erleichtern. Dies soll möglich sein, egal ob ein einzelner Pfad, mehrere oder alle Pfade eines Ensembles simultan inspiziert werden. Diese These gibt einen Einblick, wie die vorgestellten Visualisierungen entwickelt wurden, was deren Besonderheiten und Vorteile sind und welche Einschränkungen mit sich bringen. Abschließend wird darauf eingegangen, wie sich der T-Maze Explorer weiter entwickeln kann.

Abstract

behaviourists and ethologists study cognitive abilities such as learning and memory in rodents to get a better understanding of how similar processes in humans proceed. Often such studies are based on experiments of rodents placed and observed in a Multiple T-Maze. There, the path of the animals are recorded as they move inside the maze, and the resulting trajectories are then analysed. State-of-the-art analysis is based on descriptive parameters and standard statistics where one trajectory at a time is analysed. Usually it is not possible to examine multiple animal paths simultaneously. Together with experts on the field we abstracted the typical work-flow of such analyses and developed an interactive visual analytics tool, with the goal to facilitate the experts' work and enable a deeper and novel understanding of the learning ability and decision making in rodents. After giving an overview of related works and computer-aided analysis tools in the beginning, the analysis demands and task-breakdown is presented, followed by an Explanation of the data acquisition process, data preprocessing and aggregation. The underlying data structure will be explained as well. The developed analysis tool — the T-Maze Explorer — supports multiple, linked views, which support several traditional methods of visualizations, as well as two newly proposed visualizations fitted to meet the experts' analysis demands. The first view — the T-Maze View — displays all trajectories of an ensemble with additional options such as highlighting the return path. The purpose of the second view — the Gate-O-Gon view — is to extract information from the trajectories on how often returns in the path occurred and between which parts of the maze these occurred. This information is depicted in a compact and informative novel visualization. The purpose of the T-Maze Explorer is to enable its user to easily find patterns in the data and identify irregular behaviour while inspecting a single path, multiple or the whole trajectory ensemble simultaneously. This thesis provides an insight on how the proposed visualizations were developed, the T-Maze Explorer's characteristics and benefits as well as it's limitations. Lastly, a brief excerpt is given on how the T-Maze Explorer could be extended in the future.

Contents

CHAPTER 1

Introduction

Overview

In today's society people's life expectancy rises, while at the same time the population continuously increases. This is accompanied by a shift in the age distribution towards elder people over 60 and 75 years. While the population is expected to work until old age, be stress resilient, and highly active and adaptable, it is a biological fact that between the age of 50 and 65 the human organism is subjected to change, entailing a decline in physical, mental and cognitive capabilities and resources [Wei08].

Improving these capabilities — as well as combating and preventing the cognitive decline — can, among others, be achieved through improved health care. A better and deeper understanding of the complex human being's brain, neurobiology of learning and memory [Qui16] provides new pedagogical approaches and practices. New medical treatments for diverse illnesses, such as Alzheimer or autism, can be investigated.
Even nowadays a substantial amount of mysteries revolving around Memory Learning remain to be unraveled. What are the fundamental mechanics behind cognitive skills, perception and problem solving? How does the brain operate? How can cognition be modeled and intelligent systems developed? These questions belong to the field of Cognitive Science [Abd16], among which Cognitive Psychology studies memory and learning processes.

If the fundamental processes of memory and learning can be fully understood, it is possible to positively influence perception and processing of information and act accordingly. How are humans able to enact intelligent behaviour and be aware of this? How is perception of the environment and action — based on the interpretation of the received information — conducted? Utilizing memory is the act of evaluating perceived information and taking action according to previous knowledge; whereas learning changes that memory with the aim of taking improved action in the future [DDS08]. Human beings store such information in the short and long-term memory of the brain.

To verify that the process of learning actually occurred and memory was formed, it is necessary to analyse memory that was first formed in a "training" session and then recalled in a "test" sessions. behaviour based on recalling memory, which was formed in the "training" session, should give a new response when exposed to new stimuli in the same context in a "test" session [Qui16].
To conduct meaningful and scientifically correct research these analysis sessions — or experiments — have to take place in a controlled environment and under exact and reproducible conditions. Especially in regards to long-term memory experiments have to be carried out over a broad time interval. It is almost impossible to manage experiments with humans spanning years. Too many uncontrollable outside factors would not make results reproducible and letting test persons live in a controlled environment is ethically highly questionable.

It is, though, possible to study living organisms resembling humans — mammals with a short life expectancy. An eligible candidate are rodents, as their brains show significant similarities to the humans' and they tend to perceive and act similar [BN06]. Their average life expectancy is two years [Bir13]. Understanding the rodents' behaviour leads to better understanding of similar processes in humans [BLJ94]. Studies on rodents have been carried out since the beginning of the 20^{th} century [Sma01]. Typically diverse types of mazes are used to gain an understanding of learning and memory in relation to aging [PSBN12, BN06]. Commonly used mazes include the traditional "Hedge" Maze, the Single and Multiple T-Maze, its variation the Y-Maze, the Radial Arm Maze as well as the Morris Water Maze [Mor84, Ans, OD71]. This work focuses on the evaluation of results generated by a Multiple T-Maze, a maze consisting of multiple elements resembling a T, i.e. a single base and two turns — left and right. A more in-depth elaboration of this type of maze can be found in section 3.2. The tasks provided by a Multiple T-Maze are used to test left-right discrimination, references and working memory, and study how decisions evolve with different tasks and repeated runs [SHT^{+}12]. It is often used to test the effects — positive and negative — of certain substances or new medicine on the rodent's memory. While rodents explore the maze, their paths are recorded and the resulting trajectories evaluated. Commonly, these trajectories are analysed by extracting scalar features and descriptive statistics to get insight and find conclusions.

Problem Statement

While these simple features are enough to verify a positive effect on learning and memory — e.g. by testing if the goal was reached and time needed to complete a task — it does not answer questions which delve deeper into the processes happening while learning. How do rodents learn and how fast? Why do they behave in a specific way? What influences their decision making, especially correct or wrong choices?
Another aspect that can be considered is examining a single animal in contrast to a whole focus group. Of course it is possible to analyse a track in detail and gain some insight. But, if an understanding of the behaviour of an animal in a given situation can be clarified, does this apply to another animal, or even the whole group? To derive a general

understanding, the focus group must be of considerable size. Analyzing a whole ensemble of trajectories is a tedious task as multiple trajectories usually cannot be examined at the same time using traditional means except with general descriptive statistics. Without computer assistance and computer aided (pre-)processing it is difficult to understand and find patterns as well as see cause and effect relationships in data ensembles, especially regarding unstructured data. Visualizations can help in understanding data by abstracting it to a simple and easy-to-interpret visual information. They also enable the examination of different — often hidden — relationship in and between data. Embedding a computer aided system in the analysis process helps researchers to efficiently extract and compare information by providing visual guidance and feedback. This can lead to deeper and novel insight.

Contribution

Andrienko et al. refer to such a visual feedback system as "Interactive Visual Analysis" [AA13]. It enhances knowledge extraction from a large data ensemble and is applicable to tackle the above mentioned limitations of traditional analysis. This thesis describes the development of the T-Maze Explorer, an interactive visual analysis tool allowing the analysis of Multiple T-Maze trajectories from multiple perspectives and the execution of diverse analytical tasks on the whole ensemble, as identified by domain experts (see section 3.1 for details).

Spatio-temporal information is extracted from the whole trajectory ensemble, obtained through a controlled experiment, and displayed in a novel graph consisting of two main visualizations: The Multiple T-Maze view and the Gate-O-Gon view.

The **Multiple T-Maze** view shows different data abstractions, such as overall movement and possible paths, typical paths and deviations. The ensemble is represented as a CurveView displaying all trajectories at once. The view is embedded into an abstract representation of the maze used in the experiment. This gives an overview of the combined movement throughout the maze and (a-)typical behaviour of rodents. The **Gate-O-Gon** highlights the most important characteristics of the movement while concentrating on the more insight-full reverse movement. It depicts movement between T-Segments (Gates) of the maze, in particular reverse movement, and integrates a histogram of all start-end gate combinations. A circular layout combining features from a chord diagram - easy comparison of information and their internal relationships - and a flow map - distribution and location, movement and flow of data - to efficiently and compactly deliver information [Rib].

The trajectories are pre-processed and turned into representations that can be facilitated by these visualizations. The T-Maze explorer was implemented using ComVis [MFGH08] — a Coordinated Multiple View system (CMV) developed at the VRVis Research Center Vienna — as a framework. It follows Shneiderman's visual information-seeking mantra [Shn96] - overview first, zoom and filter, details on demand - to guide the user effectively through an data exploration and analysis session. The ComVis system provides a wide variety of tools developed for interactive visual analysis, used to apply different visualization techniques on one data set and displaying them simultaneously in multiple, linked views. Brushing mechanisms set run-time rules for subsets of the data and allow an easy in-depth data exploration from diverse perspectives.

Structure of this work

In this paper we further addresses methodology, technical implementation, evaluation and results, discussion and conclusion, and the future outlook of the project.

Disclaimer

The T-Maze Explorer was developed at the VRVis Research Center under the supervision of Krešimir Matković from the same company. It is based on their previous approach for interactive visual analysis of open field studies — focusing on the detailed inspection of a single trajectory. This present work extends their research to Multiple T-Maze data. The Multiple T-Maze data — consisting of approximately 400 rodent trajectories — was provided by Christiana Winding-Zavadil and Michael Balka from the University of Veterinary Medicine Vienna. The following passages give explanations and examples to illustrate the underlying theory and concepts used in the T-Maze Explorer.

CHAPTER 2

Related Work

The following sections describe works and applications which influenced the elaboration of the T-Maze Explorer. First, related works and other Visual Analysis tools specific to animal trajectories are introduced. After that a short introduction to the CMV system this work is based on is given.

2.1 Related Work

The T-Maze Explorer is influenced by several research directions. As it is an interactive system for visual analysis and includes humans in the exploration loop and evaluation, it belongs to the field of visual analytics [KKS$^+$11]. Wenqiang Cui describes Visual Analytics as the interplay between data analysis, visualization and human analytical reasoning and a sense-making loop [Cui19]. It is a field which evolved from scientific as well as information visualization. The visual analytics process can be summarized by the following steps:

1. data preparation and preprocessing
2. algorithmic analysis application
3. data presentation using appropriate visualizations
4. user derived insight
5. hypotheses testing through interaction
6. visualizations update based on user interactions

Visual analytics is applied to many commercial and research fields; some related to our work will be introduced in the following passages.

Furthermore, the T-Maze Explorer uses multiple, linked views; which are defined in the coordinated multiple views paradigm [Rob07]. More research influencing this work examine way-finding and movement data of any kind as well as animal behaviour.

Movement of any kind is a popular topic and various research has been and is being conducted on movement data. The analysis of such data harmonizes well with visual analytics. Andrienko et al. [AAB+13] did extensive research on movement data in combination with interactive visualization techniques. Applications of movement data in specific fields include sports [SJL+18], air traffic [AAFG18] or pedestrian crossings [ZFAQ13]. In particular *traffic trajectories* is a well researched topic. Andrienko et al. [AA13] examine traffic trajectories with visualization techniques and movement clustering. Chen et al. did an extended survey on state-of-the-art traffic data visualizations [CGW15]. Williams et al. study animal movement patterns using tri-axial magnetometry [WHS+17]. Jentner and Keim give an overview on how visual analytics can be used to identify patterns in traffic trajectory data [JK19].

The research on movement and way-finding tasks is also tied in psychology. Here, human way-finding is studied e.g. by Siegel and White [SW75], Montello et al. [MHRW04] and Ji-Sun et al. [KGMQ08]. When a regulated and controlled environment is needed to conduct research on behaviour and cognitive processes scientists often fall back on studying rodents, as their cognitive processes and perception seem to be similar to those of humans. Sara J. Shettleworth studies animal cognition and animal behaviour [JS01] to get a better understanding of animals' learning abilities and memory. Bubna-Littitz and Jahn examine rodents to get insight on their aging processes. In all animal related studies that are conducted in a controlled environment mazes are often used to test specific tasks and behaviour. Olton and Markowska state the use of mazes as a tool for studying conditional and place discrimination [SOM93]. Oyvind Skard studies the differences of learning between humans and rodents in a Multiple T-Maze [M.A50]. The Multiple T-Maze is — among others — a well established and defined maze [HEB+00, PJCK03]. Bubna-Littitz et al. [BLHKN81] state that the T-Maze is best mean for studying learning processes and variations can be observed a lot easier than using other methods.

As far as research reveales, no similar systems for exploring and analysing large ensembles of animal trajectories are commonly used. Other tools for analysing animal movement exist, but those primarily engage in natural, migratory movement of animals in the wild and are usually GPS based and not video tracked, opposed to typical trajectories obtained in lab-settings.
Spretke et al. propose an exploration through an enrichment approach for trajectory analysis. It integrates data enrichment features into the analysis process [SBJ+11]. Their coordinated multiple views (CMV) system uses multiple advanced visualizations — a weather interface, segment clustering windows, line charts and time series graphs — and supports CMV-typical interaction such as brushing and linking.
The work of Palleschi and Crielesi does not focus on animal movement data; its focus is to explore differences and similarities between breeding data and the various factors affecting the variability in the breeding success [PC19]. They use a CMV system showing

groupings of spatial, temporal and categorical data via tree maps, parallel coordinates or scatter plots.

The visual exploration system introduced by Li et al. is targeted to fine scale relationships of multi-species animals [LFEB20]. They focus on pair-wise or individual-to-group movement and interaction under the pretense of relatedness. To enhance the analysis process they employ both static and dynamic visualizations and employ a CMV with a timeline view, and novel views to display geo-spatial relations as well as the Relatedness chord — a chord diagram to display relatedness between multiple entities and an interactive line chart for long term pairwise relatedness view. The design and execution of the Relatedness chord is similar to the Gate-O-Gon, introduced in this paper.

While the design and layout of the Gate-O-Gon visualization is our own idea, inspiration for the circular ribbon layout is drawn from the work of Zeng et al. [ZFAQ13]. Their research on visualizing and exploring interchange patterns, emerging from large volumes of trajectory data, introduces a new visual representation, *the interchange circos diagram*. A so-called interchange pattern can be defined as the redistribution of moving objects going through a junction node in a traffic network. This can be e.g. trajectories of commuters in a subway station or a train network.
The visualization is derived from the circos figure invented by Krzywinski et al. [KSB$^+$09]. The circos figure is a wide-spread versatile visualization tool but its original purpose was the examination of mutual relationships among genomes and presenting the information in a circular ideogram layout with ribbons that connect related elements.
The interchange circos diagram is adapted from the circos figure under multiple design considerations such as visual cluttering, visual connections and providing a statistical summary. The data processing and editing, as well as the demands for the visualization, are very similar to the problems the T-Maze Explorer faces, therefore Zeng et al.'s approach is well suited for the task of analyzing animal trajectories.

2.2 TiBe and other Animal Tracking Systems

Analyzing large ensembles of animal paths is not a trivial matter and several attempts have been made to aid and simplify the analysis process. The T-Maze Explorer is primarily based on previous work of Matković et al. [MWSB12]. Their goal was to complement conventional analysis of animal paths taken from animals in an Open Field with an Object Placement Task — a test for assessing spatial reference memory and evaluating exploratory behaviour, locomotive activity and anxiety-related behaviour in rodents.

Various discrete parameters for each observation were computed for each of the 800 observations using TiBe 1.0 [Wei08], a tool for tracking animals and data acquisition often used at the University of Veterinary Medicine Vienna. The TiBe system allows the examination of one measurement in a *measurement display mode*. It shows the path and additionally a virtual raster of the field; the border area, as well as significant events. Several options are available in the *measurement display mode* to analyse the path. The

path can be combined with temporal elements to highlight the movement speed and position of an animal's path by drawing a white 'x'-marker in the path. It symbolizes the position of the animal every five seconds.
Another option for showing the movement through the Open Field is by using an animated path, similar to a short movie of the animal in the maze. For a more in-depth analysis, Matković et al. used the ComVis framework (see the section 2.3 below for details). They apply a multivariate multidimensional data model, which allows curves as single attributes in records. They use ComVis to depict all paths at once in a Curve View combined with density mapping to identify areas in the open field where multiple paths are overlapping.

Matković et al. show the efficiency of interactive visual analysis for the analysis task with a number of illustrative examples. They drill down into the data by depicting diverse aspects of the data in linked views and select sub-sets of the data by efficiently applying and combining composite brushes. They also state that this analysis process is applicable for other maze types, such as the Barnes Maze or the (multiple) T-Maze, as well.

While there are — to the best of my knowledge — no other advanced visual analytics systems dealing with animal trajectories in a maze setting, there are some tracking tools with limited (interactive, visual) analysis features. One such tool is Noldus' *EthoVision XT*. The EthoVision XT is a state-of-the-art video tracking system for recording animal activity and movement automatically ([Eth]). Its dependable tracking is versatile and adaptable to various maze types, the parameters to measure are freely choosable and has adaptable analysis settings. For the analysis task the system provides visualizations to show the trajectories with playback options. It also includes a heatmap. But it has no functionality to analyse multiple trajectories at once and has no advanced visualization except for the one mentioned.

A different system dealing with the analysis of animal movement in a maze is *BehaviorCloud*, a fully integrated cloud platform for studying animal behaviour, physiology, among others. The system can track various activities inside a maze, allows for complex behaviour coding and is adaptable for high volume or long-term tracking. It extracts usual statistical parameters from the trajectories, as well as a sequence analysis of user-defined zones, but touches more on typical data science and does not provide further visualizations except for embedding the trajectory in an image of the maze. While it provides meta data associated with a trajectory, the system provides no mechanisms to analyse multiple trajectories at once.

In conclusion, currently no interactive visual analysis system for analysing animals in a maze setting exists, which allows more in-depth analysis on multiple levels, use advanced visualizations, and allows for analyzing a trajectory ensemble. The proposed T-Maze Explorer unifies all these approaches in one system.

2.3 ComVis - A Coordinated Multiple View System

In practice, Coordinated Multiple View tools are often utilized to perform interactive visual analysis. Such a tool can be designed for a specific application domain or they can be more general. They can be mature, stand-alone tools or research tools for testing new visualization tools and paradigms and allow rapid prototyping [Rob07]. The ComVis system, developed at the VRVis Research Center, belongs to the latter. It aims to provide a platform for easy use and testing of new ideas. At the same time it aims to cover all relevant aspects for a wide variety of application domains [MFGH08]. To fulfill these demands such a system should be flexible in regards to new views and data types, supply intuitive interaction and advanced data manipulation methods, be easy to use — regardless of the user's expertise and visualization know-how — as well as to support common data types.

The system design is composed of two main components, the data manager and the list of views. The data manager is responsible for storing data in a multivariate tabular data-set as well as managing the brushes for data selection. It provides the raw data to the views. A view displays data through diverse visualization techniques and allows the creation of brushes, which are stored in the data manager. Additionally, a data view is available to show detailed, raw data to support findings from the visualizations. An analysis session can consist of multiple views, but each view accesses the same data-set.

The system allows for saving and loading of the current status of an analysis session. This feature is especially appreciated in collaborative work, where easy stop, sharing and pick-off of an analysis session is necessary. While ComVis uses its own file-format (.comvis) and binary file-format (.cvbin), it supports loading of raw data from CSV-files.

As mentioned before, an essential feature for a CMV is linking & brushing. That is, selecting and highlighting a specific area of data — effectively creating a sub-set — in one view and highlighting the same sub-set in the other linked views. In ComVis the user can use two types of brushing modes; the single brush and the composite bush. As the name suggests, the single brush mode interacts with the data using a single, adaptable brush. Composite brush mode combines any number of brushes and concatenates them using Boolean operators. A brush can be selected and adapted in size, value, position etc. at any time.

CHAPTER 3

Methodology

This section describes the data acquisition, basic concepts and technology used for the development of the T-Maze Explorer. Further, the task breakdown for an analysis session — as specified in corporation with domain experts — is explained.

3.1 Work-flow Demands & Task Breakdown

While developing the T-Maze Explorer we had a clear goal in mind: for one, to analyse and simplify the work-flow as it is carried out now by domain experts when analyzing animal trajectories in regard to behaviour. Secondly, to provide an efficient tool for complex analysis, and enabling the exploration of ambiguous questions arising during analysis sessions. The analysis tasks inter-depend on the level of analysis. These levels can be classified in a similar manner as Shneiderman's visual information-seeking mantra [Shn96].

Questions such as 'the quantity of animals which achieved the goal of the maze', 'a shift in speed and distance traveled throughout the experiment' or 'the long-term memorization state of the maze' can mostly be answered with traditional analysis. Such *high level tasks* look at the data set as a whole and raise questions concerning basic ensemble characteristics.
For *the medium level* analysis of movement patterns *identifying frequently taken paths*, *frequently traversed gate sequences* and *favoured places inside the maze* interaction and visualization can facilitate the work-flow. This intermediate analysis level poses questions regarding typical movement patterns of groups of animals which require a closer examination of the data.
At *the lowest level* the analysis focuses on a single animal or trajectory. Quantitative evaluation of the path and additional information on it — such as an animal's speed and movement orientation — are highlighted here. Tasks on this level are similar to Shneiderman's 'details on demand' and reveal detailed information of a path.

In particular, experts are interested in changes in movement direction and reverse movement. A more detailed definition and description of this concept can be found in section 3.2.

The analysis tasks can be abstracted as follows:

(i) **High Level Tasks**

- **H1** Characterize the whole ensemble by means of descriptive statistics.
- **H2** Identify possible outliers in the ensemble, not only in quantitative measures, but also at the paths level.
- **H3** Identify typical paths, especially the paths in wrong direction.
- **H4** Classify maze parts depending on traversal frequency.

(ii) **Medium Level Tasks**

- **M1** analyse sub-sets of the ensembles by applying high level tasks to reasonable sub-sets.
- **M2** analyse maze parts at different granularities (segmentation finer than gates).

(iii) **Low Level Tasks**

- **L1** Execute In-depth analysis of an individual animal.
- **L2** Examine spatial and temporal components of a single trajectory.

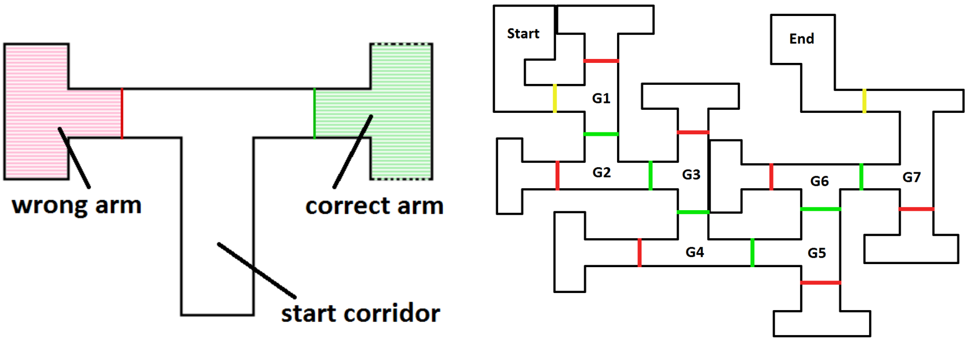
The tools provided by the ComVis framework are satisfactory for the execution of tasks **H1**, **L1** and **L2** are therefore adopted to the T-Maze Explorer. For the other tasks novel views were needed. The Multiple T-Maze view facilitates tasks **H2** and **M1**. The tasks **H3** and **M1** are supported in the Gate-O-Gon view.

3.2 Multiple T-Maze Data Acquisition

The Multiple T-Maze trajectories are provided by Christiana Winding-Zavadil and Michael Balka from the University of Veterinary Medicine Vienna. They conducted an experiment over a period of two weeks to test learning ability and memory. In this experiment several rats were placed in a Multiple T-Maze assembled inside an infrared lit room and their movements captured with an infrared camera to get their trajectories.

A T-Maze resembles a T, as can be seen in figure 3.1a. The animal starts at the base of the T-element and, moving forward, comes to a crossing where one arm is considered correct (indicated through the green area in the figure) and the other arm is classified as false (indicated through the red area). We call such a T-element *'gate'*. The goal of the single T-Maze experiment can be e.g. finding the reward — typically, food.

In a Multiple T-Mazethe correct arm leads to a new gate, i.e. the correct arm represents the base or start corridor of the next gate (indicated through the dashed lines in the green arm of figure 3.1a. This way an arbitrary amount of T-elements can be combined. The maze used by Winding-Zavadil and Balka had 7 gates and a distinct start and end area. Image 3.1b shows an the layout of the maze used by Winding-Zvadil and Balka. Each gate is labeled G_x where $x \in \{S, 1, 2, ..., 7\}$, ascending in traversing order from the start to the end area. The start area is labeled G_S instead of 0 and the end area G_E, as they are no real gates with correct/wrong arms. The first gate G_1 is considered the lowest, the last gate before the end are (G_7) the highest.



(a) (b)

Figure 3.1: The T-Maze. **a)** A single T-Maze segment. It consists of a start corridor, a crossing, the wrong arm (red) and the correct arm (green). **b)** The layout of the Multiple T-Maze used for the experiment. A start area and end area are connected through T-Maze segments G1-G7, resulting in a unique correct path.

The experiment implemented two phases: short term memory testing in the first week and long term memory testing in the second. In the first week a rat was placed inside the maze three times a day, from Monday to Friday. In the second week the same rat was tested only once on Friday. A gate is considered successfully completed if an animal crosses the border to the correct arm first. The result remains unchanged even if the animal turned around after and moved to the false arm or turned around and moved to a lower gate. To motivate the rats to explore the maze, a reward (food) was placed inside the end area. The test ended after a rat crossed the line to the end area or after a certain period of time elapsed.

By tracking the tested animals over two weeks approximately 400 trajectories were collected. The data we received from Winding-Zavadil and Balka consisted of the trajectories as a time series and already extracted simple scalar parameters, such as ’successfully finished the test’ or ’number of correctly absolved gates’, ’average time inside a gate’ and many more.

3.3 Multiple T-Maze Data Processing

To get advanced knowledge of the behaviour and cognitive abilities of rodents, the trajectories can be used as material to draw on. The position of an animal inside the maze at a given time only gives limited insight though. Therefore, the single trajectory on one hand, and the trajectory ensemble as a whole on the other have to be examined under different aspects before this newly gained information can be analysed.

The first step of this project was to extract the movement direction and direction changes from the trajectories. The animal can move along the maze in the correct — forward — direction, i.e. move through the gates in ascending order. Movement in descending order, i.e. moving from a higher to a lower gate, is deemed as false — reverse — movement. A change in direction transpires after an animal moves from a lower gate to a higher and back to a lower gate — or vice versa. More formally is a change in direction when traversal of three gates G_i, G_j and G_k occurs, where $i, j, k \in \{S, 1, 2, ..., 7\}$ and $i = k$ and $j = i \pm 1$. As one of the innate characteristics of a Multiple T-Maze is that each gate has a correct, continuing arm and a dead-end arm, it follows that a perfect, fault-free traversal exists. This perfect traversal commences in the starting area G_S, continues through all correct gates G_1 to G_7 in ascending order and finishes in the end area G_E.

The time series of a trajectory consists of a series of coordinate points corresponding to the coordinates of the Multiple T-Maze. For each coordinate point of the trajectory the associated gate was identified. Experts want to investigate where the reverse movement started and how far back an animal moved. Therefore, next, the gates where changes in direction to reverse movement — maze traversal in descending order — occurred are identified. This gate is called the start gate of the reverse movement or G_{RS}. To identify the distance traveled in the wrong direction we need to know where the following change in direction transpired, i.e. the change back from reverse to forward movement. This gate is called G_{RE}.

The gates G_{RS} and G_{RE} can now be grouped together to form a gate-pair $GP_{i,j}$, where $i, j \in \{S, 1, 2, ..., 7\}$ and $i > j$. i is the start gate G_{RS} and j the end gate G_{RE}. This is a very compact form of information on the start and end of a reverse movement sequence inside a path. For example figure 3.2 shows a possible path for the gate-pair $GP_{2,S}$. It indicates that an animal came from G_1, moved to G_2, turned around there, crossed G_1 again and finally moved back to the start area. Either the trajectory ended in G_S or the animal moved forward through the maze again.

In this manner the whole trajectory can be analysed, resulting in a set of gate-pairs. For subsequent processing this set can be represented as a binary adjacency matrix mxm, where m is the cardinal number of the gates used, in our case 8. The columns — labeled i — represent where the reverse movement started, the rows — labeled j — where reverse movement ended. The matrix is initialized as a zero matrix. For each gate-pair in the set, the respective cell in the matrix is set to true or 1. Now the information of all reverse movement of a trajectory is stored in an easy to understand, compact and versatile data structure. This matrix is called the Reverse Movement Matrix (RMM). The same process

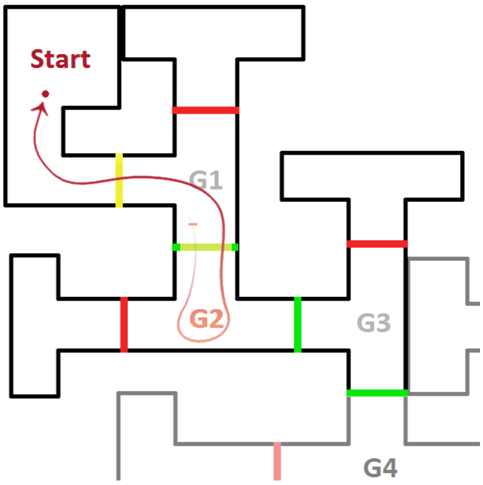


Figure 3.2: A trajectory resulting in gate-pair $GP_{2,S}$. The animal started in gate 1, turned around in gate 2 and moved back to gate S.

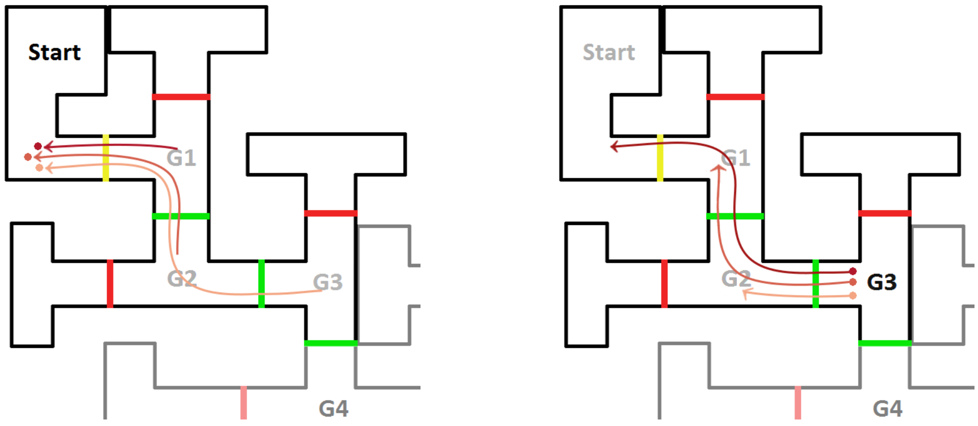
is repeated for the whole ensemble of trajectories. Instead of creating a new matrix for each trajectory, only one matrix can express the extracted ensemble information. For this, for each gate-pair occurrence the respective cell is increased by **1**. Each trajectory is associated with a unique ID. This ID is stored in a list associated with the respective cell whenever an increment occurs. This allows for precise identification of which paths are affected by certain movement patterns later on.

Figure 3.3 shows the RMM of the 400 trajectory data-set. The rows and columns are labeled $\{S, 1, 2, ..., 7\}$, i — the rows — represent the start, j — the columns — the end of a gate-pair. The value of a gate-pair can be taken from the cells. The orange highlighted cell in the RMM-matrix contains reverse movement of $GP_{3,S}$, i.e. animals moved from gate $i = 3$ to gate $j = S$ 166 times.
Understandably, the diagonal is zero as no reverse movement can occur within a single gate. The lower diagonal is symmetrical to the above one, and leads to redundant information as these values represent forward movement, except here the columns show the start and rows the end of a forward movement sequence. Values below the diagonal are zero, as this information is dispensable in our case where we are not interested in forward movement.

For better understanding, we divide reverse movement into two groups and introduce two terms to describe them: *Outgoing Reverse Movement (ORM)* and *Incoming Reverse Movement (IRM)*. Following the definition of a gate-pair, a specific gate x with $1 \leq x \leq 7$ can be the the start gate G_{RS} or end gate G_{RE} of a gate-pair. Therefore, when inspecting gate x, reverse movement can arise in two ways. First, all reverse movement starting at gate x ($x = G_{RS}$), i.e, gate-pairs of shape $GP_{x,j}$. This type of reverse movement is called *Outgoing Reverse Movement.* In the RMM, this information is encoded in the columns.

$$\text{RMM} = \begin{array}{c|cccccccc} & i=S & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline j=S & 0 & 282 & 111 & 166 & 43 & 76 & 22 & 27 \\ 1 & 0 & 0 & 34 & 26 & 6 & 10 & 4 & 4 \\ 2 & 0 & 0 & 0 & 49 & 5 & 5 & 2 & 0 \\ 3 & 0 & 0 & 0 & 0 & 9 & 5 & 3 & 4 \\ 4 & 0 & 0 & 0 & 0 & 0 & 37 & 5 & 7 \\ 5 & 0 & 0 & 0 & 0 & 0 & 0 & 25 & 7 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

Figure 3.3: The Reverse Movement matrix (RMM) compiled from the 400 trajectories. The columns, labeled j, show the distribution of Incoming Reverse Movement (IRM). The rows, labeled i, show the distribution of Outgoing Reverse Movement (ORM).



(a)

(b)

Figure 3.4: The two directions of reverse movement. **a)** Incoming Reverse Movement (IRM): all reverse movement where back-tracing ends at the same gate. Here: reverse movement stopping at gate S and coming from higher gates 1, 2 and 3. **b)** Outgoing Reverse Movement (ORM): all reverse movement where back-tracing begins at the same gate and stops at lower gates. Here: reverse movement beginning at gate 3 and ending at lower gates 2, 1 and S.

For example the column highlighted in blue in figure 3.3 shows ORM of gate 3. ORM is only possible when moving to lower gates. Figure 3.4b visualizes movement inside the maze from gate 3 to the lower gates.

Incoming Reverse Movement is reverse movement ending at gate x ($x = G_{RE}$), i.e. a gate-pair of shape $GP_{i,x}$ and can be taken from the rows of the RMM. This is highlighted in the green row in figure 3.3. Reverse movement ending at gate S can arise from gates 1 to 7. Figure 3.4a shows a cut-out of the maze and IRM to gate S, coming from gates 1, 2 and 3.

The starting point of this data analysis was the interest in gaining insight on decisions made by rodents running through the maze, with special attention to reverse movement. Therefore, the trajectory data is processed to extract reverse movement. For each trajectory, the reverse movement pairs are identified, stored and counted. The resulting reverse movement matrix can be seen as a distribution of the returns for each gate. The easiest way to display a distribution is through a Histogram. Doing so, though, would emit information on the movement flow and location. On the other hand, a visualization showing the location and movement, such as a Flow Chart or Connection Map, gives no or insufficient indication to the underlying distribution. Hence, a visualization combining orientation and location of the movement as well as the underlying distribution is desired. An additional requirement is that these specifications must be independent of the size of the data-set. With the newly developed views — described in the next chapter — we try to meet these demands.

CHAPTER 4

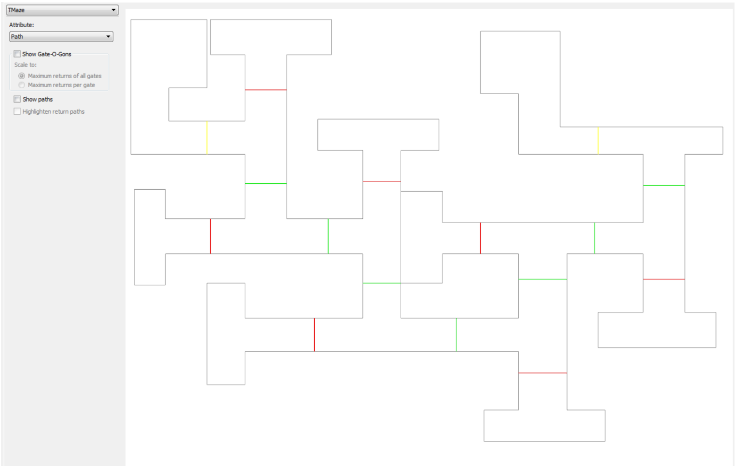
Implementation

The overall goal of the T-Maze Explorer is to help explore and analyse animal trajectories by extracting non-apparent information and displaying as much as possible in an easy to understand and useful manner. Consequentially, there is a need for a new type of visualization, as established statistical and visualization methods are not satisfactory to convey the information required for the defined tasks. Especially for data exploration, when no concrete hypotheses are formed yet, the visualization must give feedback, so that the user can adapt the analysis accordingly. This indicates that different information should be displayed at corresponding densities. Hence, we introduce two interactive charts, the T-Maze view and the Gate-O-Gon.

4.1 The T-Maze View

As the fundamental idea of this project is to enhance the analysis of animal trajectories, these trajectories should be displayed in a spatial association to the used Multiple T-Maze. Therefore, we set up a new view in the ComVis framework. This view is called the T-Maze View. Subsequently, the T-Maze View was adapted and extended to support the multiple and diverse tasks as specified before in section 3.1. After loading the Multiple T-Maze data, the outline of the used Multiple T-Maze is shown — as can be seen in figure 4.1 — including the available context menu options. In the context menu it is possible to toggle the displaying of the trajectory curves, as well as to change parameters regarding the curves.

The layout of the maze is stored as a list of T-Segments and divided into correct and wrong gates, with the start and end area explicitly defined. The trajectories are visually presented as curves and superimposed onto the T-Maze View, where the curves are adjusted to fit the maze layout (see figure 4.2a). Here, the user can visually identify outliers in the paths, as requested for **H2**. By applying a brush here or any other view, a sub-set of the data-set is selected and rendered in the color specified by the brush.



TMaze

Attribute: Path

☐ Show Gate-O-Gons

Scale to:

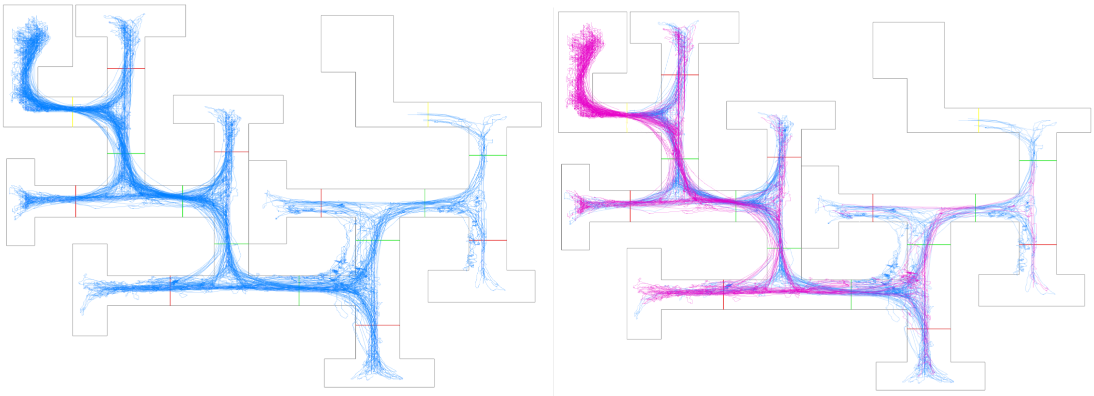
● Maximum returns of all gates

○ Maximum returns per gate

☐ Show paths

☐ Highlighten return paths

Figure 4.1: The T-Maze View with all options disabled in the context-menu and showing only the outline of the maze.



(a) (b)

Figure 4.2: Trajectories view integrated into the T-Maze View. **a)** The T-Maze View with all trajectories of the data-set superimposed. **b)** The T-Maze View with the reverse movement parts of the trajectories highlighted in purple.

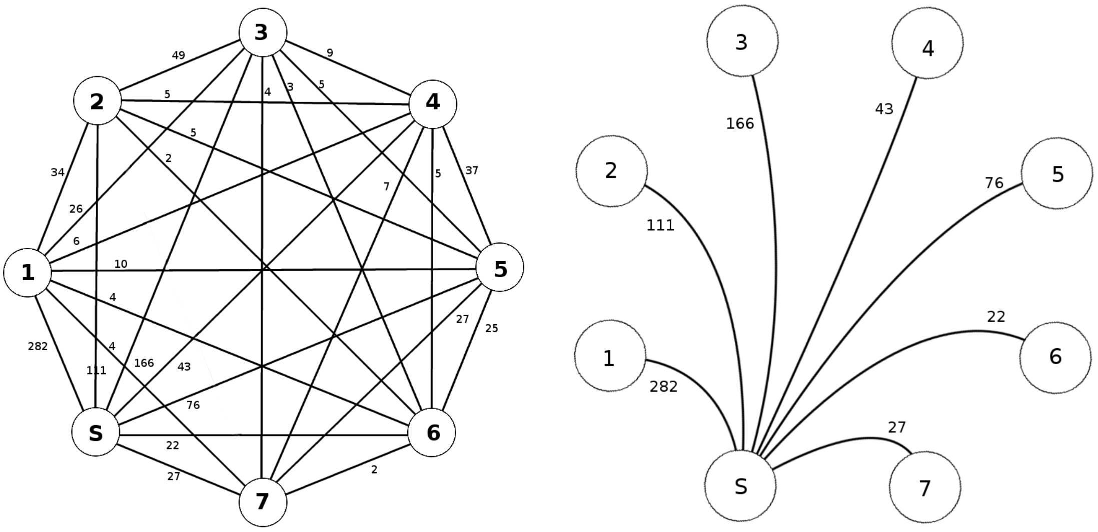
Trajectories not part of the sub-set are rendered too, but with more transparency and in light gray. This way the user retains an overview and context of the whole ensemble and can intuitively analyse sub-sets, as requested for **M1**.

In the context menu the user can choose additional options. During the data processing, the time-series are split into forward and reverse movement and labeled accordingly. It is possible to set additional visual focus on the absolute position and highlight the reverse movement. In figure 4.2b the reverse movement parts are painted in purple.

4.2 Gate-O-Gon

4.2.1 Design Evolution

Until the final design, the Gate-O-Gon went through several phases to adjust to diverse requirements and problems. Based on the fact that the RMM is an extended adjacency matrix, the first design was a simple adjacency graph shown in 4.3a as a simple adjacency graph of the trajectory ensemble data-set. The gates are represented as nodes in the graph. An existing gate-pair is indicated by a line between the two corresponding nodes, with the quantity of occurrences displayed next to the line. Nodes are arranged circularly and clockwise in ascending order. While it gives an overview of the existing gate-pairs, it impedes quick information acces and conveys neither the distribution nor the movement direction intuitively.



(a) (b)

Figure 4.3: First visualizations of the RMM adjacency matrix. **a)** The RMM displayed as a simple adjacency graph. Nodes represent the gates, lines the gate-pairs. **b)** The adjacency graph after being split by gate, forming a tree. Here: Adjacency-Tree of gate S.

To counter the overloading, we split the graph into a distinct adjacency tree for each gate. Again, the nodes represent the gate, but only the connections between the currently inspected gate — where incoming reverse movement occurred — and higher gates is shown. The node of the current gate is considered the root of the tree. By splitting one visualization showing all connections into 8 separate graphs, the connections become more apparent. The position of the nodes are static, which makes it easier to compare different adjacency trees. In figure 4.3b, the IRM from higher gates to the start gate are displayed in an adjacency tree. In this case, it is obvious that from each higher gate several animals returned to the start. Still, information on the quantity of such returns is missing.

Consequently, the next step was to include the distributions. As stated before, the easiest way to represent the distribution of returns from each gate is via a histogram. When inspecting a single gate, domain experts are more interested in the higher gates where animals returned to the currently inspected one, i.e. the incoming reverse movement. As described in section 3.3, this information is encoded in the rows of the RMM.
Generating the IRM histograms for each gate, these histograms can now be superimposed on the nodes of the adjacency graph. After some revisions on the simple histogram design, the appearance of the nodes in the adjacency tree was swapped with the histogram-nodes as shown in figure 4.4a.
While it might be sufficient to only show the histogram of the current gate in the node, it is more difficult to compare the distributions of different gates. By exchanging every node with the respective histogram, the user can maintain an overview of all distributions at any time. To enhance the focus on the currently inspected gate and its the affiliated histogram, only the current gate node is rendered in full color, while the other histograms are only hinted at in gray, to provide context (more on the color used in the Gate-O-Gon is explained in section 4.2.2).

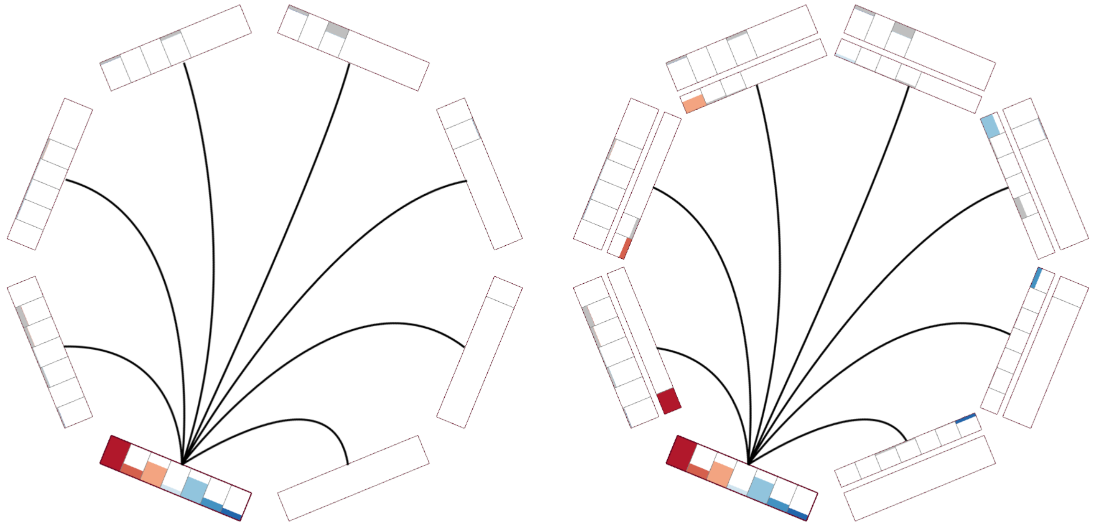

(a) (b)

Figure 4.4: Integration of reverse movement distribution as histograms. **a)** The resulting graph for gate S after incorporating the IRM distributions. The nodes representing the gate are replaced by the respective distribution. **b)** By stacking the outgoing and incoming reverse movement distributions, more information can be displayed.

Though there is higher preference on IRM, it is desirable to also depict the ORM distributions. Again, histograms can be used to show the distribution. The histograms are integrated into the nodes, stacked on top of the IRM histograms (figure 4.4b). The ORM histograms are scaled down to give emphasis to the IRM histograms of each node. As each bin of the histogram relates to a specific gate, emphasis is set on the bin associated with the currently inspected gate through highlighting it in full color, same as the IRM histogram of the current gate. The other bins provide context but don't draw the focus on them by depicting them in gray.

By providing the ORM, the user not only gains insight on where animals came from to the current gate, but is also provided with information on where other outgoing reverse movement of a specific gate ended. For example, when examining the start gate in figure 4.4b, it is apparent that a lot of animals returned from gate 5 (big, light blue bin). When looking at the ORM histogram of gate 5, we see that most animals moved back to the start gate S. Approximately half as many returned to gate 4 but almost none to any other gates. This information is less evident when only the IRM is viewed.

Now, comparing the different reverse movement distributions is straightforward. It is evident at first glance where reverse movement occured by examining the connection-lines between the gate-nodes. Now it is possible to execute high level tasks satisfactorily. But we still want to enhance the insight gained when observing all gate-graphs at once. Placing the 8 graphs side-by-side, the user can grasp from where to where animals moved by examining the connection-lines. While the distributions are visible, the effect can be heightened by integrating the IRM distributions into the lines, similar to a chord diagram. Chord diagrams are used to compare similarities within a data-set or between different groups of data ([Rib] - Chord Diagram).

The inter-relationship between the gates (IRM gate-pairs) are displayed using arcs consisting Bézier curves, i.e. an arc is displayed between the ORM histogram of a certain node and the IRM of the currently observed gate if they share a gate-pair (i.e. the value of the related cell in the RMM is greater zero). One end of the arc is docked to the ORM histogram of the start gate i of the gate-pair $GP_{i,j}$. The other end docks to the bin associated with the gate i in the IRM histogram of the current gate j. Figure 4.5 depicts the graph for gate S with nodes connected through arcs.
The distribution of the IRM is integrated into the thickness of an arc proportionally to the quantity of the returns of this gate-pair. Consequently, the user can now examine the trajectory ensemble in an overview and clearly perceive where reverse movement occurred and its extend compared to other gates. This final graph of reverse movement, as depicted in figure 4.5, is called Gate-O-Gon. Each gate-segment is represented through its own Gate-O-Gon.

Lastly, to emphasize the spatial relationship of a gate and a Gate-O-Gon, the Gate-O-Gons are embedded into the T-Maze view and each gate is superimposed on its associated gate-segment. The visibility of the Gate-O-Gons can be set through the context menu in the T-Maze view. Figure 4.7 shows all Gate-O-Gons superimposed on the Multiple T-Maze. The Gate-O-Gon associated with the gate S is placed in the middle of the start area of the maze and is called Gate-O-Gon S. The Gate-O-Gon 1 is placed in the middle of the correct segment of gate 1. And so on. The size of the Gate-O-Gon corresponds to

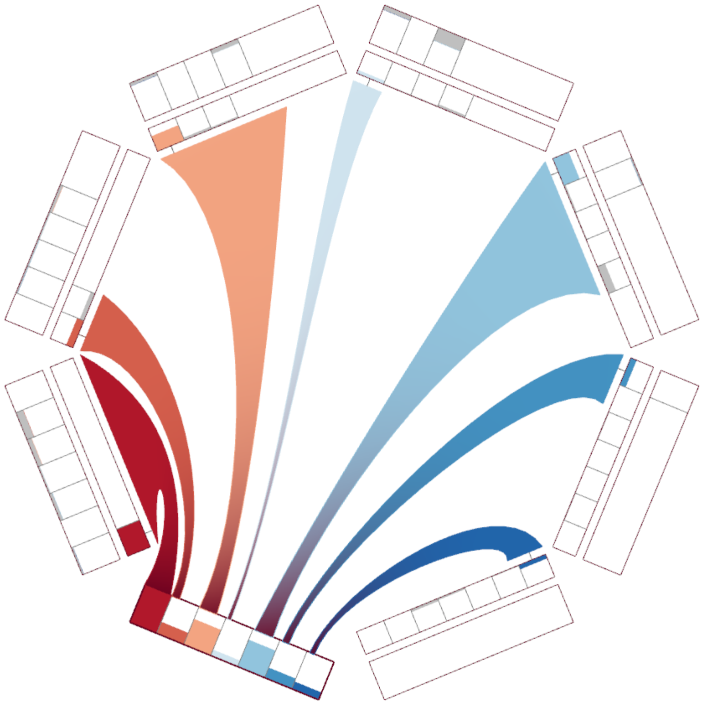


Figure 4.5: The final design of the Gate-O-Gon of gate S. Arcs, with width reflecting the IRM and ORM distribution, enhance the inter-relationship between gates and allow easy overview comparison of reverse movement of different gates.

the display dimensions of the T-Maze view and scales automatically when the view is resized. The Gate-O-Gon size adapts to the minimal distance between all Gate-O-Gons. Doing so, the maximum size of a Gate-O-Gon without overlapping one another is ensured.

4.2.2 Gate-O-Gon Color Scheme

To better distinguish the gates internally and to augment the relation of gate-pairs, each gate is associated with a distinct color. Color is often used to distinguish different categories, in this case the different gates. Associating a distinct color to a specific gate facilitates making comparisons and detecting patterns. This can help executing task **H3** - finding typical or atypical paths in the data-set. Further, to augment the affiliation of the gate and its color, the Gate-O-Gon outline is also rendered in a shade of the respective color.

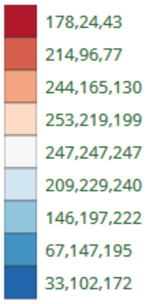


Figure 4.6: The diverging color scheme for 9 classes. Taken from ColorBrewer [BHS+].

The used colors are interpolated between dark red, light blue and dark blue. This

color scheme is provided by *ColorBrewer* [BHS$^+$], a diagnostic tool for evaluating the robustness of individual color schemes. The chosen color scheme is of diverging nature, color-blind safe and was chosen with nine distinguishable classes in mind. Figure 4.6 shows the 9 distinct color classes.
While there are other schemes that meet these criteria, we chose the red-blue one, as humans tend to associate red with a negative outcome and blue with a more neutral one. In our case the color red is associated with gate S, as this is the furthest point where animals can move back to and consequently the 'worst' outcome. The further along the T-segment sequence an animal returned, the less negative the outcome, and therefore the gate color turns from red to blue.

By using a diverging color scheme the visualization becomes aesthetically more pleasing, which in general is appreciated by users, which in return makes working with the visualization more enjoyable.

4.2.3 Gate-O-Gon Scaling

In a typical assumed analysis session a user will examine the data by means of multiple diverse views. The user might be interested in an overview to get a feeling for the particularities of the data. In some cases more detailed information might be desired. In others, putting the data into a different context may lead to new insight. This is true for the Gate-O-Gon as well. Therefore, diverse scaling modes are appropriate; on a spatial level and on a data context level.

The first spatial level is the overview of all Gate-O-Gon's, also called the Gate-O-Gon Overview, as seen in figure 4.7. Here, the user can analyse the whole data set in the context of all Gate-O-Gon's and their placement in the maze, as is also the first level of Shneiderman's mantra. However, the histograms might be very small due to restricted display or view size.
The second level is inspecting a single Gate-O-Gon. By clicking on the requested Gate-O-Gon, the system switches to the Single Gate-O-Gon mode. In figure 4.8a the maze in the background is hidden, as well as the other Gate-O-Gon's. The selected Gate-O-Gon is zoomed to fit the size of the view. The user can better examine the reverse movements related to the selected Gate-O-Gon. Additionally to the enlarged size, supplementary information to the reverse movement distributions is given. In the left upper corner the name of the selected Gate-O-Gon is stated. In the case of figure 4.8a: *'Gate-O-Gon: S'*. Each gate node is labeled with the respective gate number.
Lastly, the absolute values of the IRM and ORM distributions are displayed to give the user concrete insight into the data.

In the context menu, the user can choose to scale the data according to *'maximum returns of all gates'* or *'maximum returns per gate'*. *'maximum returns of all gates'* is the default scaling mode and is used in figure 4.8a. Here the scaling of the histogram bins is global and proportional to the maximum value in RMM. The reverse movement that occurred the most is taken as 100% and the other bins are scaled in relation to this value. Figure

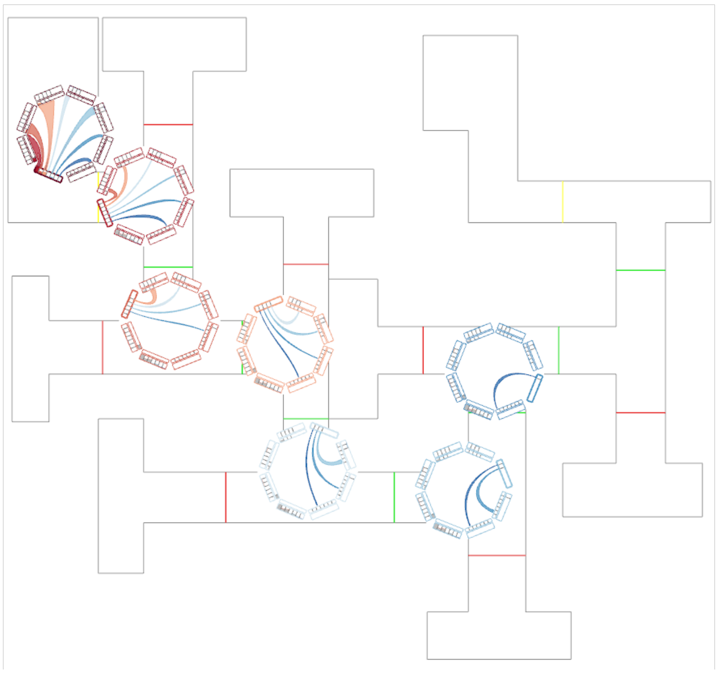
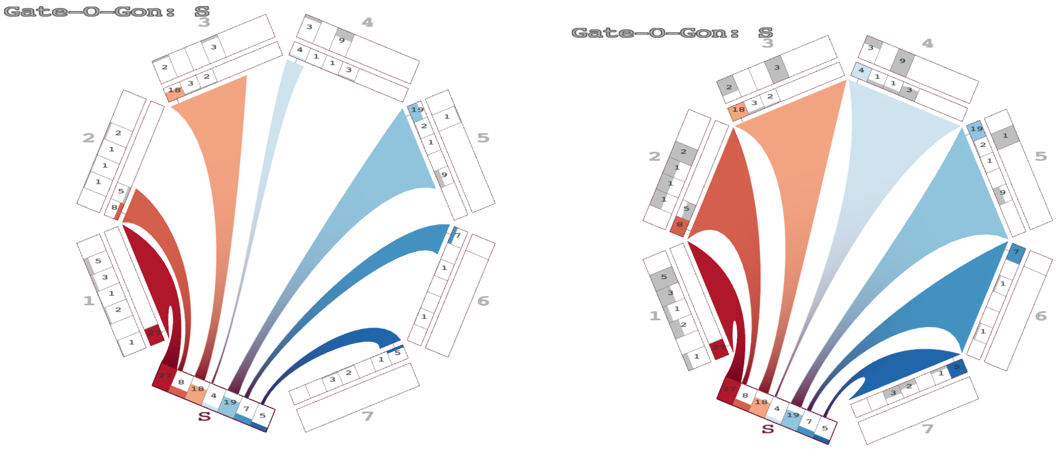


Figure 4.7: The T-Maze View with paths hidden and Gate-O-Gons shown in 'Overview' mode.



(a)

(b)

Figure 4.8: The two scaling modes of the Gate-O-Gon. **a)** The detail-view of the Gate-O-Gon ***S*** (Start) in standard scaling mode *'maximum returns of all gates'*, additionally displaying the absolute values of the reverse movement in the histograms. **b)** The detail-view of the Gate-O-Gon ***S*** when scaled relative to *'maximum returns per gate'*.

4.8a shows that in the IRM histogram of the start gate, the first bin — returns from gate 1 to gate S — is completely filled, indicating that the maximum reverse movement happened between gate-pair $G_{1,S}$. This scaling mode highlights the maximum movement without hiding small numbers.

The second scaling mode, *'maximum returns per gate'*, considers each gate separately. Thus, scaling happens proportional to the maximum returns (IRM) of a gate each. Figure 4.8b shows the Gate-O-Gon S when scaled per gate. As the arc's width is scaled according to the IRM and ORM value at their respective ends, it is evident in this scaling mode, that for each gate, the highest outgoing returns are reverse movement to the gate S. This information on maximum returns is emphasized in this scaling mode.

4.2.4 Brushing

As mentioned before, brushing is an essential feature for a Coordinated Multiple View system. In the T-Maze View brushing is available in the Single Gate-O-Gon mode. Any Gate-O-Gon can be selected to apply a brush. A new brush is created as soon as the user clicks on a specific non-empty bin in any of the histograms or an IRM arc. Internally, the brush stores the ID of the row and column of the RMM cell associated with the selected histogram bin. Therefore, the brush is not affected by which Gate-O-Gon was selected and is globally applied. The brush selects a sub-set of trajectories which contain reverse movement between the selected gate-pair. The brush can either be a Single Brush or Composite Brush with the same Boolean concatenations applying to it as in the other ComVis views.

In figure 4.9 *Top* a Single Brush is used to select a sub-set of the trajectories containing the gate-pair $GP_{4,S}$. The whole data-set is still visible but with decreased opacity. The selected sub-set is highlighted in the arcs, IRM and ORM and its thickness and height respectively are scaled relatively from the value of the sub-set to the value of the whole data-set. Figure 4.9 shows that none of the animals, which have a gate-pair $GP_{4,S}$ have a gate-pair $GP_{6,S}$. Approximately half of the animals returning from gate 1 also returned from gate 4 go gate S. Data which was brushed in one of the linked views is reflected in the T-Maze View as well. Figure 4.9-Bottom shows the subset of the data-set after brushing the gate-pair $GP_{4,S}$ in the Gate-O-Gon.

4.3 Informal Case-Study

The T-Maze explorer is a system developed for domain experts to facilitate their work. It is a successful enrichment for the researchers if the work-flow can be improved and they gain new insights through the use of our tool. In the scope of this thesis, no formal case-study has been conducted yet by domain experts. However we carried out a first small, informal case-study in the scope of a conference paper for the *'Eurographics Workshop on Visual Computing for Biology and Medicine'* ([BSM18]). For this case-study, we followed the work-flow of an analysis session and tried to perform several tasks. The case-study reveals that the exploration phase of the analysis is enhanced and increases efficiency in performing the tasks. During the case-study many interesting and new findings were revealed, which warrants the validity of the T-Maze Explorer. First feedback from domain experts substantiate these claims.

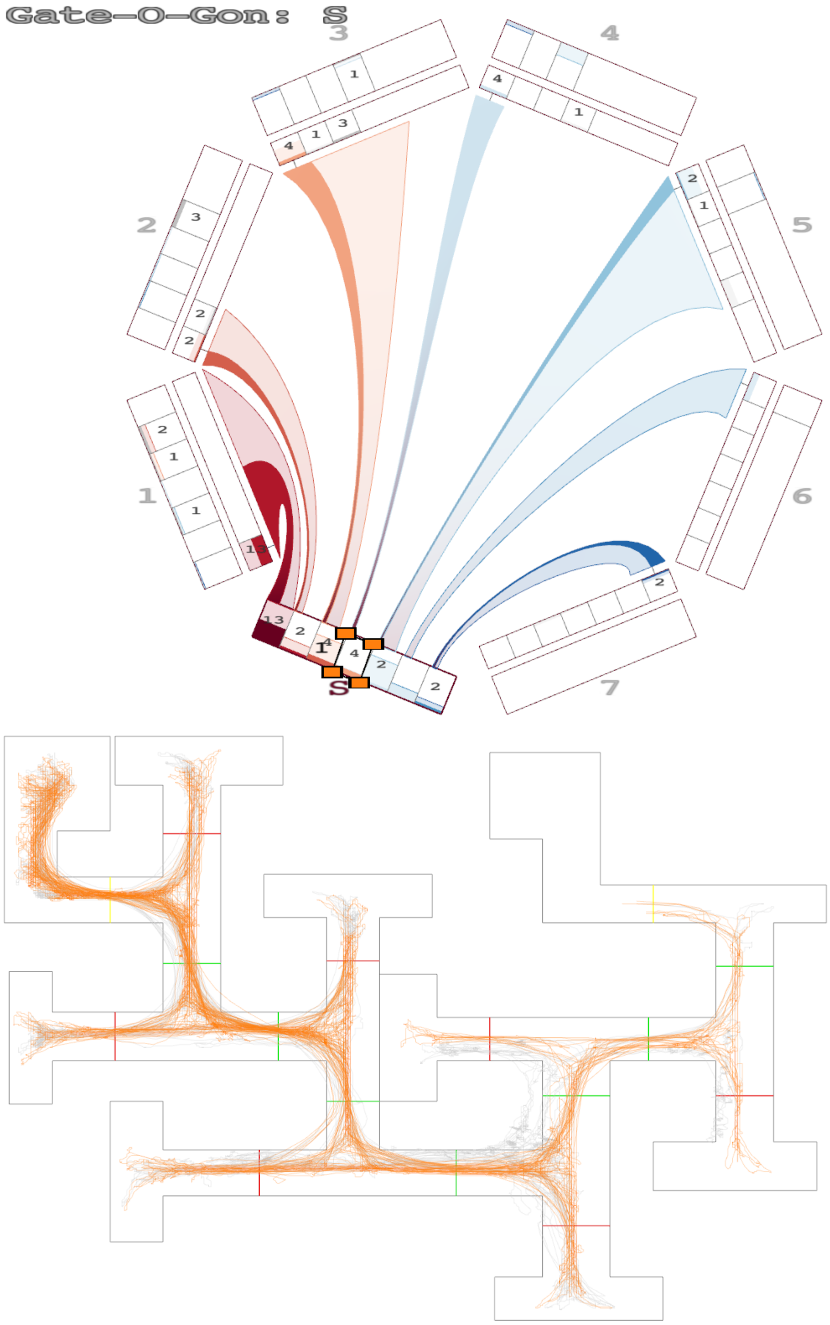


Figure 4.9: *Top:* The detail-view of the Gate-O-Gon *Start* with a brush applied to the bin of the IRM histogram associated with gate-pair $GP_{4,S}$, effectively selecting only trajectories containing reverse movement from gate 4 to gate S. *Bottom:* The T-Maze View highlighting the trajectories containing reverse movements from gate 4 to gate S.

CHAPTER 5

Conclusion & Discussion

This thesis describes a new approach for analyzing ensembles of animal trajectories in a Multiple T-Maze. In cooperation with domain experts we abstracted a typical analysis session and specified a list of tasks, which shall be facilitated through interactive visualizations. We explain how the trajectories are analysed to extract information on the reverse movement and why it is important to differentiate between incoming and outgoing reverse movement in relation to a specific area of the maze. The design and implementation of the T-Maze Explorer, a new interactive visualization system we developed to support the analysis sessions, is explained1.

The T-Maze View provides an environment to examine trajectory paths in the context of the used Multiple T-Maze with special attention to reverse movement and outliers on the path level. The Gate-O-Gon is a new interactive visualization, developed to analyse the reverse movement of all animals throughout the maze. We give insight into the benefits of the design and how encountered problems were tackled. Different insight can be gained from the Gate-O-Gon, depending on the data context it is placed in. Lastly we show how selecting sub-sets of the data through brushing can strengthen the knowledge-gain and also lead insight. While the introduced system has many advantages, it is not yet formally approved by domain experts. Therefore, further improvements and extensions on the T-Maze Explorer to fully support domain experts in their work-flow and facilitate their performance is desired.

CHAPTER 6

Extensions & Future Work

While this visualization opens new possibilities for exploring Multiple T-Maze trajectory data, there is still room for improvement. At the current state, the low level tasks **L1** and **L2** are not covered within the T-Maze Explorer, as the old TiBe system [MWSB12] satisfactorily fulfills the requirements for these tasks. Future work could include the extension of the available features in the TiBe system and integrating them into the T-Maze Explorer to unite all functionality in one system. The T-Maze Explorer was extended outside of the scope of this thesis to meet the demands of the tasks not tackled here and to improve others. To analyse the popularity of parts of the maze at different levels a multi-resolution Heatmap was introduced ([BSM18]). The Heatmap allows the examination of the data on a high level to see which gates are highly frequented and which are not (**H4**), but also to drill down to a pixel level to detect highly frequented spots in the maze.

Domain experts have also shown interest in analyzing the trajectories as a sequence of traversed gates in combination with posing search queries on these sequences. This improves the performance of high level tasks as it can ease the detection of outliers and patterns. Another approach to facilitate the analysis is through automatic or semi-automatic analysis. Based on the traversal sequence, automatic analysis can complement the detection of clusters in the data-set. A first approach using different clustering metrics and text comparison metrics has already been integrated ([BASM19]). A combination of automatic analysis and a visual feedback to refine the results can lead to more insightful results.

List of Figures

Bibliography

[AA13] Natalia Andrienko and Gennady Andrienko. Visual analytics of movement: An overview of methods, tools and procedures. *Information Visualization*, 12:3–24, 01 2013.

[AAB+13] Gennady Andrienko, Natalia Andrienko, Peter Bak, Daniel Keim, and Stefan Wrobel. *Visual Analytics of Movement*. 12 2013.

[AAFG18] G. Andrienko, N. Andrienko, G. Fuchs, and J. M. C. Garcia. Clustering trajectories by relevant parts for air traffic analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):34–44, Jan 2018.

[Abd16] Asad Abdi. Cognitive science; why cognitive science? -, 09 2016.

[Ans] Anne Anson. "rats and mazes" rat behaviour and biology. `http://www.ratbehavior.org/RatsAndMazes.htm`. Accessed: 27.06.2019.

[BASM19] Fabrizia Bechtold, Hrvoje Abraham, Rainer Splechtna, and Krešimir Matkovic. Interactive Pattern Analysis of Multiple T-Maze Data. In Tatiana von Landesberger and Cagatay Turkay, editors, *EuroVis Workshop on Visual Analytics (EuroVA)*. The Eurographics Association, 2019.

[BHS+] Cynthia Brewer, Mark Harrower, Ben Sheesley, Andy Woodruff, and David Heyman. Colorbrewer: Color advice for cartography. `http://colorbrewer2.org/`. Accessed: 15.10.2019.

[Bir13] J.E. Birren. *Handbook of the Psychology of Aging*. Handbooks of aging. Elsevier Science, 2013.

[BLJ94] H. Bubna-Littitz and J. Jahn. Psychometric testing in rats during normal ageing. procedures and results. *J Neural Transm Suppl*, 44:97–109, 1994.

[BN06] Jennifer L. Bizon and Michelle M. Nicolle. 32 - rat models of age-related cognitive decline. In P. Michael Conn, editor, *Handbook of Models for Human Aging*, pages 379 – 391. Academic Press, Burlington, 2006.

[BSM18] Fabrizia Bechtold, Rainer Splechtna, and Kresimir Matkovic. Visual exploratory analysis for multiple t-maze studies. In Anna Puig Puig, Thomas Schultz, Anna Vilanova, Ingrid Hotz, Barbora Kozlikova, and Pere-Pau Vázquez, editors, *Eurographics Workshop on Visual Computing for Biology and Medicine*, pages 203–213. The Eurographics Association, 2018.

[CGW15] Wei Chen, Fangzhou Guo, and Fei-Yue Wang. A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems*, 16:2970–2984, 12 2015.

[Cui19] Wenqiang Cui. Visual analytics: A comprehensive overview. *IEEE Access*, pages 1–1, 06 2019.

[DDS08] Nathaniel D. Daw and Daphna Shohamy. The cognitive neuroscience of motivation and learning. *Social Cognition*, 26:593–620, 10 2008.

[Eth]

[HEB+00] Harald Hoeger, Mario Engelmann, Guenther Bernert, Rainer Seidl, Hermann Bubna-Littitz, Wilhelm Mosgoeller, Barbara Lubec, and Gert Lubec. Long term neurological and behavioral effects of graded perinatal asphyxia in the rat. *Life Sciences*, 66(10):947 – 962, 2000.

[JK19] Wolfgang Jentner and Daniel Keim. *Visualization and Visual Analytic Techniques for Patterns*, pages 303–337. 01 2019.

[JS01] Sara J. Shettleworth. Animal cognition and animal behaviour. *Animal Behaviour*, 61:277–286, 02 2001.

[KGMQ08] Ji-Sun Kim, Denis Gračanin, Krešimir Matković, and Francis Quek. Finger walking in place (fwip): A traveling technique in virtual environments. In Andreas Butz, Brian Fisher, Antonio Krüger, Patrick Olivier, and Marc Christie, editors, *Smart Graphics*, pages 58–69, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[KKS+11] Joern Kohlhammer, Daniel A. Keim, Giuseppe Santucci, Gennady Andrienko, and M. Pohl. Solving problems with visual analytics. In *The European Future Technologies Conference and Exhibition 2011*. Procedia Computer Science, 2011.

[KSB+09] Martin I Krzywinski, Jacqueline E Schein, Inanc Birol, Joseph Connors, Randy Gascoyne, Doug Horsman, Steven J Jones, and Marco A Marra. Circos: An information aesthetic for comparative genomics. *Genome Research*, 2009.

[LFEB20] Wei Li, Mathias Funk, Jasper Eikelboom, and Aarnout Brombacher. Visual exploration of movement relatedness for multi-species ecology analysis. 01 2020.

[M.A50] Oyvind Skard M.A. A comparison of human and animal learning in the stone multiple t-maze. *Acta Psychologica*, 7:89 – 109, 1950.

[MFGH08] Kresimir Matkovic, Wolfgang Freiler, Denis Gracanin, and Helwig Hauser. Comvis: a coordinated multiple views system for prototyping new visualization technology. In *Proceedings of the 12th International Conference Information Visualisation*, pages

[MHRW04] Daniel R. Montello, Mary Hegarty, Anthony E. Richardson, and David Waller. chapter Spatial Memory of Real Environments, Virtual Environments, and Maps., pages 251–285. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, 2004.

[Mor84] Richard Morris. Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of neuroscience methods*, 11(1):47–60, 1984.

[MWSB12] Kresimir Matkovic, Christiana Winding, Rainer Splechtna, and Michael Balka. Interactive visual analysis of ethological studies: Getting insight from large ensembles of animals' paths. In *EuroVA 2012: International Workshop on Visual Analytics*, pages 85–89, 2012.

[OD71] John O'Keefe and Jonathan Dostrovsky. The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 1971.

[PC19] Alessia Palleschi and Matteo Crielesi. A visual analytics system of data gathered from colonial seabirds. pages 224–227, 07 2019.

[PJCK03] Panaree Parnpiansil, Nuanchan Jutapakdeegul, Thyon Chentanez, and Naiphinich Kotchabhakdi. Exercise during pregnancy increases hippocampal brain-derived neurotrophic factor mrna expression and spatial learning in neonatal rat pup. *Neuroscience Letters*, 352(1):45 – 48, 2003.

[PSBN12] Calvin P. Stone and Dorothy Bird Nyswander. The reliability of rat learning scores from the multiple-t maze as determined by four different methods. *The Pedagogical Seminary and Journal of Genetic Psychology*, 34:497–524, 09 2012.

[Qui16] Jorge Alberto Quillfeldt. *Behavioral Methods to Study Learning and Memory in Rats*, pages 271–311. Springer International Publishing, Cham, 2016.

[Rib] Severino Ribecca. The data visualisation catalogue. Accessed: 17.08.2019.

[Rob07] Jonathan C. Roberts. State of the Art: Coordinated & Multiple Views in Exploratory Visualization. In *Proc. of the 5th International Conference on Coordinated & Multiple Views in Exploratory Visualization.* IEEE CS Press, 2007.

[SBJ+11] David Spretke, Peter Bak, Halldor Janetzko, Bart Kranstauber, Florian Mansmann, and Sarah Davidson. Exploration through enrichment: A visual analytics approach for animal movement. pages 421–424, 11 2011.

[Shn96] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE Symposium on Visual Languages*, pages 336–343, Sep 1996.

[SHT+12] Hirotaka Shoji, Hideo Hagihara, Keizo Takao, Satoko Hattori, and Tsuyoshi Miyakawa. T-maze forced alternation and left-right discrimination tasks for assessing working and reference memory in mice. *Journal of visualized experiments : JoVE*, 60, 02 2012.

[SJL+18] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. Andrienko, M. Grossniklaus, and D. A. Keim. Bring it to the pitch: Combining video and movement data to enhance team sport analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):13–22, Jan 2018.

[Sma01] Willard S. Small. Experimental study of the mental processes of the rat. ii. *The American Journal of Psychology*, 12(2):206–239, 1901.

[SOM93] David S. Olton and Alicja Markowska. Chapter 8. mazes: their use in delayed conditional discriminations and place discriminations. *Techniques in the Behavioral and Neural Sciences*, 10, 12 1993.

[SW75] Alexander W. Siegel and Sheldon H. White. The development of spatial representations of large-scale environments. volume 10 of *Advances in Child Development and Behavior*, pages 9 – 55. JAI, 1975.

[Wei08] Silvia Weis. *Effekte kognitiver Stimulierung auf alternde Sprague-Dawley Ratten (Tiermodell "Pensionsschock").* PhD thesis, Veterinärmedizinische Universität Wien, 2008.

[WHS+17] Hannah Williams, Mark Holton, Emily Shepard, Nicola Largey, Brad Norman, Peter Ryan, Olivier Duriez, Michael Scantlebury, Flavio Quintana, Elizabeth Magowan, Nikki Marks, Abdulaziz Alagaili, Nigel Bennett, and Rory Wilson. Identification of animal movement patterns using tri-axial magnetometry. *Movement Ecology*, 5, 03 2017.

[ZFAQ13] Wei Zeng, Chi-Wing Fu, Stefan Müller Arisona, and Huamin Qu. Visualizing Interchange Patterns in Massive Movement Data. *Computer Graphics Forum*, 2013.